Modelling Cause-of-Death Mortality and the Impact of Cause-Elimination

Daniel H. Alai [1] Séverine Arnold (-Gaille) [2] Michael Sherris[3]

School of Mathematics, Statistics and Actuarial Science
University of Kent, Canterbury, Kent CT2 7NF, UK

Department of Actuarial Science
Faculty of Business and Economics (HEC Lausanne)
University of Lausanne, 1015 Lausanne, Switzerland

CEPAR, Risk and Actuarial Studies, Australian School of Business
UNSW, Sydney NSW 2052, Australia

Abstract

The analysis of causal mortality provides rich insight into changes in mortality trends that are hidden in population level data. Therefore, we develop and apply a multinomial logistic framework to model causal mortality. We use internationally classified cause-of-death categories and data obtained from the World Health Organization. Inherent dependence amongst the competing causes is accounted for in the framework, which also allows us to investigate the effects of improvements in, or the elimination of, cause-specific mortality. This has applications to scenario-based forecasting often used to assess the impact of changes in mortality. The multinomial model is shown to be more conservative than commonly used approaches based on the force of mortality. We use the model to demonstrate the impact of cause-elimination on aggregate mortality using residual life expectancy and apply the model to a French case study.

Keywords: Cause-of-Death Mortality, Multinomial Logistic Regression, Cause-Elimination, Life Expectancy, Mortality Forecasts

JEL Classifications: G22, G32, C51, C18

[1] d.h.alai@kent.ac.uk
[2] severine.arnold@unil.ch
[3] m.sherris@unsw.edu.au

1 Introduction

Mortality rate improvements have significantly impacted life insurance as well as private and public pension systems. Cause-of-death data provides additional insight into these improvement trends. We add to the area of causal mortality modelling by developing and applying the multinomial logistic model. This allows a better understanding of the potential impact of the various causes of death on aggregate mortality. The multinomial logistic model provides a framework for cause-elimination that is easy to employ and complements current methods used in practice. It offers a new perspective on the potential impact of medical innovation. It has applications in life insurance for insurers that are developing mortality scenarios, and is particularly relevant with recent developments in Solvency 2 regulations. It can also be applied to investigate the impact of possible future mortality trends in current European pension reforms.

The multinomial *logistic* model (also known as multinomial *logit* model) is typically used to detect factors that significantly influence a polytomous response; that is, a response with several competing outcomes. Several applications of the multinomial logistic model have been undertaken with respect to cause-of-death analysis over the past ten to twenty years. Examples include Eberstein *et al.* (1990), who used eight categorical and continuous independent variables, including marital status, education, and birth weight, to model five infant cause-specific mortality rates. Lawn *et al.* (2006) applied the multinomial logistic framework to model the distribution of neonatal deaths in countries with poor data; see Johnson *et al.* (2010) and Liu *et al.* (2012) for related work. Bradshaw *et al.* (2003) and Shahraz *et al.* (2012) employed a multinomial model to redistribute unknown or ill-defined deaths; see Murray *et al.* (2006) for a related application to ill-defined causes. Park *et al.* (2006) incorporated the multinomial logistic framework in the modelling process in order to take into account the impact of the tenth revision of the international classification of diseases.

However, none of these studies have investigated cause-specific mortality over the entire age-range. Previous applications have been solely on infant mortality. As mentioned by Foreman *et al.* (2012), *current techniques do not allow for us to take advantage of such modelling advances within a multinomial framework.* Since past studies were mainly interested to find the variables that have the largest impact on cause-specific mortality, many variables were included in the model, and thus computational power issues arose. However, this is not an issue when relatively few regressors are included in the model. Our focus is on utilizing a framework that accounts for the nature of competing risks, that provides additional insights to existing research based on joint risk factors, frailties, hazard rates, and copulas. Since cause-specific mortality data typically includes two variables of interest, namely age and time, the multinomial logistic model is easy to employ. More importantly, the multinomial logistic framework parsimoniously quantifies the impact in the event that a cause is (partially) eliminated; for example, in case a cure is found for a specified disease. This significantly increases the applicability of the model framework.

The competing nature of cause-of-death mortality rates has challenged many researchers over the past. Indeed, the study of competing risks brings necessarily the interesting question of dependence. To put our work in perspective, we briefly

review various other cause-specific mortality studies that have been interested in the dependence/independence assumptions over the past few decades. The regularly-used independence assumption is among the limitations of cause-specific mortality studies. For a broader discussion on the benefits, limitations and risks of decomposing mortality by cause of death, see e.g. Tuljapurkar (1998), Gutterman and Vanderhoof (1998), Tabeau *et al.* (2001), Booth and Tickle (2008) and Richards (2009).

The benefits of forecasting mortality for each cause in isolation and then aggregating them to forecast total mortality rates has been an area of interest. Such studies assume that the causes of death are independent and, as such, can be forecast independently from one another. For example, McNown and Rogers (1992) used univariate ARIMA models to forecast the parameters of a multi-exponential function fitted to the age pattern of mortality. Based on data from 1960 to 1975, they forecast four main causes of death (heart diseases, cancer, vascular diseases, accident and violence) until 1985. Similar studies include Caselli (1996) and Wilmoth (1996), who considered the impact on projections of modelling mortality rates by cause; Rogers and Gard (1991), who illustrated several applications of the Heligman-Pollard function, one of them used to forecast cause-specific mortality; Wilmoth (1995), who demonstrated that for specific models, such as the Lee-Carter model, overall mortality forecasts were consistently lower than the sum of mortality forecasts based on a cause-specific approach; and Caselli *et al.* (2006) and Tabeau *et al.* (1999), who compared several forecasting approaches applied to aggregate as well as cause-specific mortality rates.

Various models have been developed that attempt to account for the dependence between causes. For example, cause-specific mortality rates have been correlated through joint dependence on individual risk factors (covariates); see e.g. Rosén (2006) and Manton (1986). Frailties have also been widely employed to account for heterogeneous populations, where the dependence assumptions between the various causes are determined by the joint distribution of the frailties; see e.g. Manton *et al.* (1986), Vaupel and Yashin (1983) and Hougaard (1984). Multiple cause-of-death data allow the investigation of links between various causes and help to determine a *pattern of failure*, defined as a combination of causes that result in death; see e.g. Manton *et al.* (1976); Manton and Poss (1979); Manton *et al.* (1980a); Manton and Myers (1987). More recently, copulas have been used to model the dependence between competing risks. Kaishev *et al.* (2007) investigated (partial) cause-elimination by extending the double-decrement results of Carriere (1994) to include up to four causes. Dimitrova *et al.* (2013) generalized the copula approach to include cause-elimination by *ignoring* and *eliminating* causes using the definition of these terms introduced in Elandt-Johnson (1976). See Lo and Wilke (2010) for related work.

However, in practice cause dependence is not generally taken into account. Using individual risk factors or multiple causes requires significant additional data that may not be readily available, whilst the frailty model and the copula framework are more complex and thus less convenient to apply. Therefore, the most widely used approach is still based on a model developed more than 40 years ago by Chiang (1968), in which causal forces of mortality are used and which provides insight into causal trends and partially addresses dependency issues; see e.g. Prentice *et al.* (1978). For example, this approach has been widely used in past cause-elimination

and cause-delay models; see e.g. Keyfitz (1977), Tsai *et al.* (1978), Manton *et al.* (1980b), Olshansky (1987, 1988), and Manton (1991), amongst others. Since 1968, the United States decennial life tables have been published with a special report that focuses on the impact of eliminating causes using Chiang's approach (which will also be referred to as the *force of mortality approach* or the *instantaneous approach*); see Bayo (1968), Greville *et al.* (1975), Curtin and Armstrong (1988), Anderson (1999). United States official projections and forecasts of the Institute of Actuaries of Australia are both performed under the force of mortality approach; see Wong-Fupuy and Haberman (2004) and LIWMPC Longevity Research Group (2010), respectively.

This paper aims to provide a soundly based, and relatively easy to apply, alternative to complement Chiang's traditional approach. For that purpose, the multinomial logistic model is a natural choice since causal mortality data is necessarily polytomous in nature. Furthermore, the simplicity of the model specification will allow for the incorporation of a diverse range of cause-specific mortality dependence. We assume a simple and intuitive relationship between the cause-of-death mortality rates, based on a probabilistic approach. The model provides a convenient tool for cause-elimination studies and scenarios analysis. To emphasize the complementary nature of the two approaches, we provide a comparison between the multinomial approach combined with probabilistic point of view (further referred only as the *multinomial approach*) and the method developed by Chiang.

After introducing the methodology with respect to the multinomial logistic model and life expectancy calculation in Section 2, the probabilistic approach is presented in Section 3. Total and partial cause-eliminations are introduced by shocking causal mortality. It is shown that survival increases comparatively more in the force of mortality approach than in the multinomial logistic approach. In Section 4, we illustrate the model in a case study using data for France obtained from the World Health Organization. Section 5 concludes the paper.

2 Methodology

In this section we provide the theoretical details of the proposed causal mortality model. We also outline the construction of residual life expectancy using an abridged life table.

2.1 Multinomial Logistic Model

Multinomial logistic regression techniques are catered to modelling probabilistic response variables for competing outcome categories; see e.g. Menard (2002) and Borooah (2002). The model describes a polytomous response by using a sequence of binary responses. For example, survival or death is the first binary response; if death, death by cancer or not, could be the second binary response, etc.

Let $D_i(x,t)$ denote the random deaths from cause i for age x at time t and let $L(x,t)$ denote the subsequent survivors that complement the deaths. Consider n causes and define $Y(x,t)$ to be the vector of cause-specific deaths and survival. We have

$$Y(x,t) = (D_1(x,t), D_2(x,t), \ldots, D_n(x,t), L(x,t))'.$$

We assume $Y(x,t)$ follows a multinomial distribution, whose probability mass function, omitting the arguments (x,t), is given by

$$\Pr[D_1 = d_1, \dots D_n = d_n, L = l] = \frac{E!}{d_1! \cdots d_n! l!} q_1^{d_1} \cdots q_n^{d_n} p^l,$$

where,

$$\sum_{k=1}^{n} q_k(x,t) + p(x,t) = 1,$$

such that $q_k(x,t)$ describes the probability of death as a result of cause k, $p(x,t)$ the probability of survival, and

$$E(x,t) = l(x,t) + \sum_{k=1}^{n} d_k(x,t),$$

where $l(x,t), d(x,t)$ are realizations of the random variables $L(x,t), D(x,t)$, and the resulting measure of initial exposure is given by $E(x,t)$. We adopt survival as the baseline category in the multinomial logistic framework. The baseline-category logit model produces the following:

$$\log \frac{q_i(x,t)}{p(x,t)} = X(x,t)\beta_i, \quad i = 1, \dots, n,$$

where $X(x,t)$ is the design matrix and β_i the vector of regression parameters especially suited to cause i. The design matrix, $X(x,t)$, contains values of explanatory variables; these may be indicator or numerical variables for categorical or continuous covariates, respectively. The product of the design matrix and the vector of regression parameters is called the linear predictor, or the regression formula, which we outline in Section 2.2 below. Given the regression parameters and the design matrix, we apply the logistic function to obtain the probabilities of interest, which are given as follows:

$$q_i(x,t) = \frac{\exp\{X(x,t)\beta_i\}}{1 + \sum_k \exp\{X(x,t)\beta_k\}}, \quad i = 1, \dots, n, \tag{1}$$

$$p(x,t) = \frac{1}{1 + \sum_k \exp\{X(x,t)\beta_k\}}. \tag{2}$$

Notice that the form of the survival probability, $p(x,t)$, differs from the probabilities of death, $q_i(x,t)$, since survival is designated as the baseline category.

2.2 The Regression Formula

Given the multinomial framework, we address the structure of the regression formula. The regression links our response to any potential covariates and is typically some combination of age, period, and cohort. There is a vast literature which investigates this component of aggregate mortality modelling starting with the seminal work of Lee and Carter (1992); in addition, some overviews are provided by e.g. Cairns *et al.* (2011) and Haberman and Renshaw (2011).

The nature of our data suggests the exclusion of any overt cohort covariate. The practical reason is twofold. First, we generally have a limited number of periods,

which hinders our ability to identify any significant cohort trend. Second, cause-specific data are presented in age-groups; we comment further on this impact in our case study below. Given that we have data by age-group, in order for cohort effects to be taken into account, the data would need to be converted to reflect single age covariates, which is a non-trivial exercise in itself. There is also an overriding theoretical reason why we avoid cohort considerations. Namely, causes of death have an intuitive relationship with periodic developments, particularly those due to medical innovations.

Whether a covariate should be treated as categorical or continuous is a second point of consideration. Categorical covariates offer more flexibility but can over-burden the model. We consider categorical *age* and continuous *period* covariates. Categorical age is both intuitive and convenient. Intuitive, since it is likely that the various age-groups exhibit contrasting behaviour with respect to the different causes of death. Convenient, since we have a limited number of age-groups. Likewise, continuous period is both intuitive and convenient. Intuitive, since mortality over time is typically classified as a *trend* whose underlying behaviour is of a functional form. Convenient, since implementing continuous time avoids resorting to time-series analysis for forecasting purposes. Lastly, to treat *both* age and period as categorical would be most flexible, but would also be susceptible to overfitting. The number of parameters in such a model would approach the number of observations. This would result in a model with near-perfect fit, but lacking the ability to produce sensible prediction.

Finally, it has been observed in the literature that various age-groups react differently to time; see e.g. Booth *et al.* (2001). Therefore, we allow for age-period interaction. The linear regression formula we adopt is as follows:

$$\eta_i(x,t) = \beta_{0,i} + \beta_{1,i,x} + f(t;\tilde{\beta}_{i,x}),$$

where

$$\eta_i(x,t) = \log \frac{q_i(x,t)}{p(x,t)},$$

and $f(t;\tilde{\beta}_{i,x})$ is a function defining the age-period interaction. Note that the linear regression parameters are distinct for each cause i. The subscript x on $\beta_{1,i}$ and $\tilde{\beta}_i$ indicates the relevant age-group and the tilde on $\tilde{\beta}_i$ signifies it is a vector of parameters. Parameters are estimated using maximum likelihood. Several environments for statistical computing contains functions fitting the multinomial logit model on some dataset; see, for example, the *mlogit* package in *R* or the *logistic* procedure in *SAS*.

2.3 Residual Life Expectancy

In order to present easily-interpretable outcomes we use (residual) life expectancy. Since we work with age-groups rather than individual ages, we make use of the abridged life table method; see e.g. Chiang (1984). This method mirrors that of a standard life table, with some modifications to allow for the interval age-groups. It requires an assumption on the relationship between central and crude mortality rates governed by a parameter denoted a_x. This parameter takes the interpretation of the average proportion of the year lived for those that died. Throughout the

paper, we assume $a_x \equiv 0.5$, which results in the following relationship between q, the crude, and m, the central mortality rate,

$$q(x,t) = \frac{2m(x,t)}{2 + m(x,t)};$$

see Section 4.1 for more details. Even if this is an assumption that could be challenged for infant mortality, it is widely used and accepted for adult age mortality, which is the focus of this paper.

3 Causal Mortality Shocks

A particular interest in the field of mortality concerns the impact of medical innovation in the form of cures and any corresponding increase in longevity; the cure for cancer being a particularly prevalent example. In contrast to the study of aggregate mortality, a causal approach provides the framework in which valuable insight can be gained.

In this section, we outline how causal mortality is shocked in the multinomial logistic model, and compare it with the approach based on modelling forces of mortality. By causal mortality shock, we mean that mortality for a specific cause suddenly increases, decreases or is eliminated due to some event, such as an epidemic or the discovery of a new cure. The remaining cause-specific mortality rates are subsequently also affected by the change applied to the shocked mortality.

3.1 Shocks in the Multinomial Model

First, we acknowledge the possibility that the elimination of a cause can initiate a marked increase in some causes, whilst decreasing or not affecting others. However, any such relationship is, strictly speaking, unobservable. To understand these particular relationships is a non-trivial matter and is not the aim of this paper. Without the consideration of these causal relationships, the description of cause-elimination in the model is more faithfully represented by the idea of *ignoring* causes, rather than *eliminating* them, based on the definition of these terms introduced in Elandt-Johnson (1976).

For completeness, we briefly describe these definitions. Let T_i be the time of death from cause i alone. With n potential causes of death, consider a joint distribution for the random vector $T_{i=1,\ldots,n}$. It is clear that we are able only to observe one of these times of death, that is, $T = \min(T_1, T_2, \ldots T_n)$. Elandt-Johnson (1976) considered *elimination* as a conditional limiting distribution where time of death by the eliminated cause approaches infinity; that is, $T_i \to \infty$ when cause i is eliminated. Besides, $T = \min(T_1, T_2, \ldots, T_{i-1}, T_{i+1}, \ldots, T_n)$ is considered as *ignoring* T_i, i.e. *ignoring* cause i. The two notions are equivalent under the assumption of independent times of death T_i. Most of the existing literature operates under the notion of *ignoring* a cause; a notable exception is the recent work of Dimitrova *et al.* (2013). Henceforth, we continue to use the term cause-elimination, but this should not be confused with the definition outlined in Elandt-Johnson (1976).

In this paper, we approach the problem from a probabilistic point of view with no prior knowledge. Hence, if one of the competing outcome categories is eliminated, we assign its probability proportionally to the other outcomes, where survival is

merely one of these outcomes. That is, although survival probability will certainly increase as a result of a cure for cancer, it will not do so on a one-to-one basis with the decrease in cancer-specific mortality. This mechanism resembles the independent causal forces-of-mortality assumption used in instantaneous modelling approaches; this is demonstrated in Section 4.3 in a comparison with the results of Kaishev *et al.* (2007).

Suppose we introduce a shock $\rho_{i,x} \geq 0$ to cause i and age-group x, where values of $\rho_{i,x} > 1$ signify a marginal increase in mortality, and vice versa. Note that $\rho_{i,x} = 0$ corresponds to the elimination of deaths by cause i for age-group x. The resulting probabilities are adjusted as follows:

$$q_i(x,t) = \frac{\rho_{i,x} \exp\{X(x,t)\beta_i\}}{1 + \sum_k \rho_{k,x} \exp\{X(x,t)\beta_k\}}, \quad i = 1, \ldots, n, \tag{3}$$

$$p(x,t) = \frac{1}{1 + \sum_k \rho_{k,x} \exp\{X(x,t)\beta_k\}}. \tag{4}$$

We also introduce a shock ρ_i that uniformly affects all age-groups; that is, a shock independent of age-group. We continue by using and applying the assumption of age-independent shocks, however, the theory is not made more complex by allowing age-dependent shocks $\rho_{i,x}$. Equations 3-4 become:

$$q_i(x,t) = \frac{\rho_i \exp\{X(x,t)\beta_i\}}{1 + \sum_k \rho_k \exp\{X(x,t)\beta_k\}}, \quad i = 1, \ldots, n,$$

$$p(x,t) = \frac{1}{1 + \sum_k \rho_k \exp\{X(x,t)\beta_k\}}.$$

We provide a brief example to clarify the impact of cause-elimination without prior knowledge; that is, under the proportional reweighting mechanism. Consider exposure to causes of death 1 and 2 for some specific age x and year t, with probabilities 1/3 and 1/6, respectively. Consequently, the survival probability is 1/2. Using equations 1-2, we have:

$$q_1(x,t) = \frac{\exp\{X(x,t)\beta_1\}}{1 + \exp\{X(x,t)\beta_1\} + \exp\{X(x,t)\beta_2\}} = \frac{2/3}{1 + 2/3 + 1/3} = 1/3,$$

$$q_2(x,t) = \frac{\exp\{X(x,t)\beta_2\}}{1 + \exp\{X(x,t)\beta_1\} + \exp\{X(x,t)\beta_2\}} = \frac{1/3}{1 + 2/3 + 1/3} = 1/6,$$

$$p(x,t) = \frac{1}{1 + \exp\{X(x,t)\beta_1\} + \exp\{X(x,t)\beta_2\}} = \frac{1}{1 + 2/3 + 1/3} = 1/2.$$

The elimination of cause 1 would have the following impact on the remaining two probabilities:

$$q_2(x,t) = \frac{1 * 1/3}{1 + 0 * 2/3 + 1 * 1/3} = \frac{1}{4},$$

$$p(x,t) = \frac{1}{1 + 0 * 2/3 + 1 * 1/3} = \frac{3}{4}.$$

The above is a consequence of applying equations 3-4 together with the conditions $\rho_1 = 0$ and $\rho_2 = 1$. Alternatively, the probabilities may equivalently be formulated

as

$$\begin{aligned} q_2(x,t) &= \frac{1}{6} + \frac{1}{3}\,\frac{1/6}{1/6+1/2} = \frac{1}{4}, \\ p(x,t) &= \frac{1}{2} + \frac{1}{3}\,\frac{1/2}{1/6+1/2} = \frac{3}{4}. \end{aligned}$$

That is, each probability is increased by a proportion of the eliminated mass 1/3. The proportions are determined by the weights of the respective probabilities prior to elimination. This is not the consequence of the multinomial logistic model, but rather the consequence of the elimination *mechanism* we chose, based on the probabilistic approach. This mechanism is the canonical approach when prior knowledge of relationships amongst the causes is unknown. However, with such knowledge, different forms of dependence may be incorporated into the model. Furthermore, the desired adjustment mechanism is independent of software packages or functions and may be suitably customized. Revisiting the above example, if it is known that the elimination of cause 1 shifts probability entirely to cause 2 (leaving the survival probability unchanged), this would result in $q_2 = p = 1/2$. Using equations 3-4, this is obtained with conditions $\rho_1 = 0$ and $\rho_2 = 3$. Therefore, jointly specifying the shock parameters ρ_i allows for the implementation of many forms of dependence.

It is evident that we adjust for mortality shocks on an *annual* probability basis, while the traditional method of Chiang (1968) adjusts the causal force of mortality, which is representative of the instantaneous probability of death by cause.

Consider the survival probability as written in terms of the force of mortality:

$$p(x,t) = \exp\left[-\int_0^1 \mu(x+s,t)ds\right],$$

where $\mu(x,t) = \sum_k \mu^{(k)}(x,t)$. That is, the total force of mortality, $\mu(x,t)$, is the addition of the forces of mortality attributed to each cause. The effects of causal mortality shocks are imposed by shocking the appropriate component of the force of mortality. For example, cause j elimination is achieved by removing the relevant component of the total force of mortality and subsequently recalculating the survival probability; resulting in:

$$p(x,t) = \exp\left[-\int_0^1 \sum_{k\neq j} \mu^{(k)}(x+s,t)ds\right].$$

Compared with our annual approach, probability redistribution on an instantaneous basis favours survival. In other words, when cause j is eliminated in our method, deaths from causes $i \neq j$ increase comparatively more and survival increases comparatively less than previous findings that modelled causal forces of mortality. A formal proof is provided below.

3.2 A Comparison of Annual and Instantaneous Mortality

In this section we compare the impact of cause-elimination on the survival probability under the annual approach (based on the multinomial logistic model) and the instantaneous approach (based on force of mortality modelling). We show that under cause-elimination, the instantaneous approach increases survival comparatively more than the annual approach.

Given the force of mortality, a survival probability may be written as

$$p(x,t) = \exp\left[-\int_0^1 \sum_k \mu^{(k)}(x+s,t)ds\right] = \prod_i p'_i(x,t),$$

where $p'_j(x,t)$ is the *net* survival probability for cause j,

$$p'_j(x,t) = \exp\left[-\int_0^1 \mu^{(j)}(x+s,t)ds\right].$$

The net survival probability is interpreted as the survival probability if no causes of death other than cause j exist, as opposed to the *crude* survival probability, $p_j(x,t) = 1 - q_j(x,t)$, that competes with other causes. In the instantaneous approach, the elimination of cause j results in

$$p^{(-j)}(x,t) = \exp\left[-\int_0^1 \sum_{k\neq j} \mu^{(k)}(x+s,t)ds\right] = p(x,t)/p'_j(x,t),$$

where the superscript $(-j)$ in $p^{(-j)}(x,t)$ indicates the elimination of cause j. Under the constant force of mortality assumption,

$$\mu((x+\delta),(t+\tau)) = \mu(x,t), \qquad 0 \leq \delta, \tau < 1,$$

the net survival probability for cause j is known to be,

$$p'_j(x,t) = p(x,t)^{q_j(x,t)/q(x,t)};$$

see e.g. Bowers *et al.* (1986) for a proof. Thus, to find the new survival probability when cause j is eliminated, one has to divide the current survival probability by $p(x,t)^{q_j(x,t)/q(x,t)}$.

In contrast, the elimination of cause j in the annual approach that employs the multinomial logistic model results in a survival probability given by

$$p^{(-j)}(x,t) = p(x,t) \cdot \left[1 + \frac{q_j(x,t)}{p(x,t) + \sum_{k\neq j} q_k(x,t)}\right].$$

Given that both approaches result in a proportional effect on the annual survival probability, we investigate the relation between these two proportions. That is, we show that

$$\frac{1}{p(x,t)^{q_j(x,t)/q(x,t)}} > 1 + \frac{q_j(x,t)}{p(x,t) + \sum_{k\neq j} q_k(x,t)}.$$

By applying some simple algebra, we find the above inequality by proving the following:

$$(1 - q_j(x,t))^{q(x,t)} \quad > \quad p(x,t)^{q_j(x,t)}. \tag{5}$$

Inequality (5) is proved by using Newton's generalized binomial theorem and by noting that $0 < q_j(x,t) < q(x,t) < 1$; see Appendix A for a detailed proof. This shows that under cause-elimination, the instantaneous approach increases survival comparatively more than the annual approach.

4 Case Study

4.1 Data

The World Health Organization (WHO) maintains a comprehensive cause-of-death mortality database (World Health Organization (2012)). This database provides the mid-year population and number of deaths by cause for various countries over the last 50 to 60 years. We obtained data for France from 1952 to 2008. The data is generally divided into five-year age-groups with a final group for ages 85 and above. We consider France due to its size and influence in Europe.

To ensure consistency across countries, the WHO database classifies the causes according to the International Classification of Diseases (ICD); see Table 1. Under the ICD, the underlying cause of death is specified as *the disease or injury which initiated the train of morbid events leading directly to death, or the circumstances of the accident or violence which produced the fatal injury.* We consider the five main ICD causes, which are: diseases of the circulatory system, cancer, diseases of the respiratory system, external causes, and infectious and parasitic diseases. The major causes accounted for more than 80% of deaths in recent years, and made up approximately 60% – 70% of deaths 50 years ago. The cause classification used throughout the paper is summarized in Table 2.

Table 1: International Classification of Diseases - Coding System

Causes of death	ICD 7	ICD 8	ICD 9	ICD 10
Circulatory system	A079-A086	A080-A088	B25-B30	I00-I99
Cancer	A044-A060	A045-A061	B08-B17	C00-D48
Respiratory system	A087-A097	A089-A096	B31-B32	J00-J99
External causes	A138-A150	A138-A150	B47-B56	V00-Y89
Infectious and parasitic diseases	A001-A043	A001-A044	B01-B07	A00-B99

Table 2: Cause-of-Death Codification

Cause	Code
Infectious and parasitic diseases	1
Cancer	2
Circulatory system	3
Respiratory system	4
External causes	5
Other	6

Some adjustments are made in order to analyze data consistently over time. First, the number of deaths of unknown age are distributed proportionally across the age range, as recommended by the Human Mortality Database (Human Mortality Database (2012)). The proportional distribution of the number of deaths of unknown age is the classical method used by demographers and has a very limited impact in our analysis. Indeed, since 1993, there is no death of unknown age, while in the 1950's, the proportions of these deaths were usually lower than 0.05% for each age-group and cause.

Second, due to a large difference in the nature of infant mortality, the first five-year group is split into two. Thus, our database is composed of nineteen groups, the first for infants less than one year old, a second for children aged one to four, thereafter in groups of five years, ending with the group aged 85 and above. It is clear that the presence of age-groups results in a loss of information, perhaps the age-group of 85 and above being the most striking example. However, the loss of information is proportional only to the interaction between age and cause in each particular group. For example, if there are no significant relative changes amongst the causal mortality rates over the age-group, there would be no loss of information. Clearly, we anticipate some interaction in the 85 and above age-group, but not of the magnitude that would render the results meaningless.

Third, the data contains *central* exposure-to-risk rather than *initial* exposure-to-risk. Consequently, the ratio of cause-specific deaths to exposure produce *central* death rates $m_i(x,t)$ for cause i; see e.g. Pitacco *et al.* (2009) (Ch. 2) for an overview of basic mortality models. Central death rates are typically assumed to relate to death probabilities as follows:

$$q(x,t) = \frac{m(x,t)}{1 + (1 - a_x)m(x,t)}.$$

As mentioned in Section 2.3, we define $a_x \equiv 0.5$ and obtain the relationship

$$q(x,t) = \frac{2m(x,t)}{2 + m(x,t)}.$$

Finally, an adjustment is necessary due to the changes of classification over time. Indeed, the ICD changed three times between 1950 and 2010, from ICD-7 to ICD-10. This was done in order to account for progress in science and technology and to achieve more refined cause descriptions. Consequently, the raw data are not directly comparable over time. To make them comparable, comparability ratios are used.

At the time of a change in classification, some countries recorded the cause of death according to the previous classification as well as the newly adopted one. This *double death registration* makes it possible to analyze the impact of a change of classification. Unfortunately, many countries did not apply this approach for all causes. That is, they recorded deaths under both classifications for a subset of the data. Some countries did not even apply it for a single cause. We develop our own comparability ratios in order to smooth the death rates across the classifications, as in Gaille and Sherris (2011). This approach facilitates a consistent analysis across countries should such a comparison be of interest.

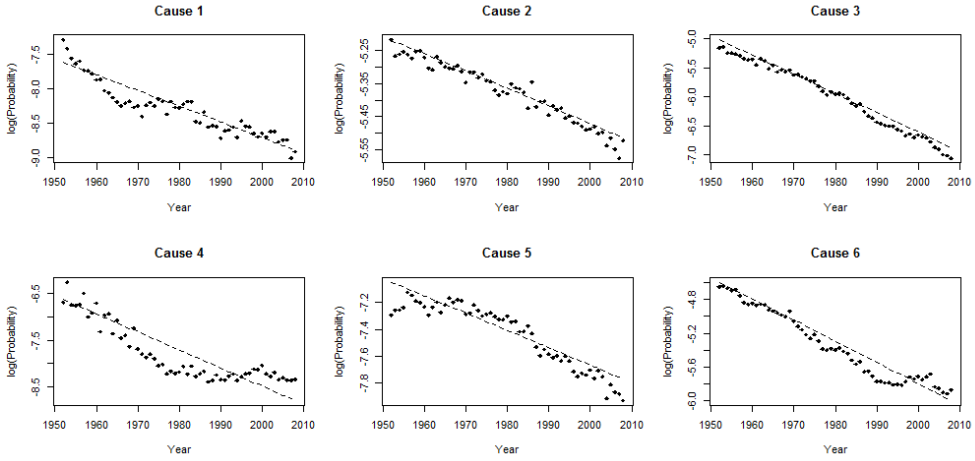
Gaille and Sherris (2011) define a comparability ratio by requiring the average of the death rates over the last two years of a classification to coincide with the average of the death rates over the first two years of the newly adopted classification. Since France adopted ICD 8 in 1968, ICD 9 in 1979 and ICD 10 in 2000, three sets of comparability ratios are developed. Comparable data over the complete period under consideration are obtained by dividing the number of deaths in a new classification by the comparability ratio connecting this classification with the previous one, etc. This ensures that mortality rates are continuous at the junction points between classifications. The following analysis uses these adjusted death rates for women in France.

4.2 Model Fitting

We begin by studying the observed mortality rates. Figure 1a presents the log-mortality rates over time for the age-group 65–69; and Figure 1b presents them over age-group for calendar year 2008. We opt to display this age-group and calendar year since they are most relevant to retirement systems and most recent, respectively.

Figure 1: Observed (log) Mortality Rates for Women in France

(a) Over time (age 65–69)



(b) Over age-group (calendar year 2008)

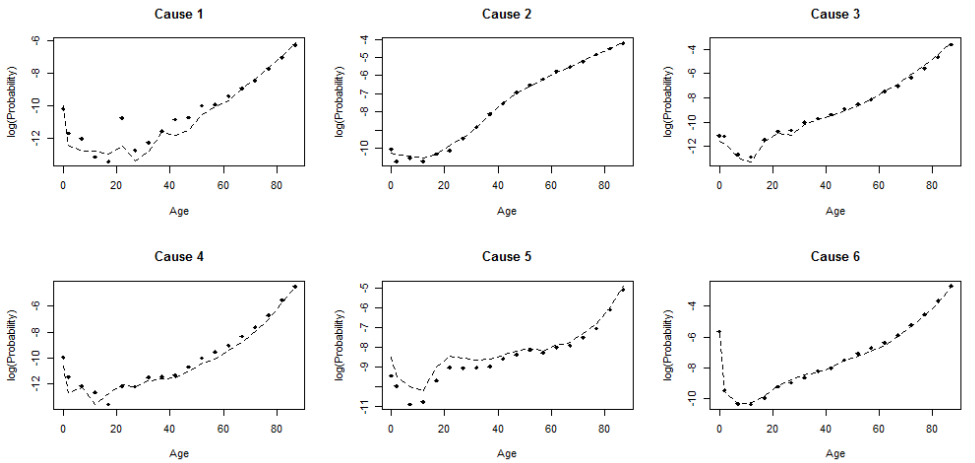


Figure 1a suggests minor quadratic behaviour, however, a linear time component appears sufficient to capture the period trend. The plots over age-group exhibit the various familiar components of the average log-mortality age pattern; see e.g. Heligman and Pollard (1980). We adopt the following regression formula for women in France:

$$\eta_i(x,t) = \beta_{0,i} + \beta_{1,i,x} + \beta_{2,i,x} \cdot t.$$

The resulting fit is presented in Figures 1a and 1b with dashed lines. The dataset contains $7,581$ observations and is fit using 228 parameters. The number of parameters required in the model is the product of the number of causes, the number of

age-specific parameters, and the number of parameters required for the functional form of time. In our example it is $6 \times 19 \times 2 = 228$. With the number of causes, no parameter reduction can be obtained. We have 19 age-categories and therefore, 19 age-specific parameters. However, age need not be specified as categorical, and we would certainly not advocate this if single age data were available. In such cases, the age dimension may be reduced, for example, to eight or nine parameters using a Heligman-Pollard functional; Heligman and Pollard (1980). The parameters required for the functional form of time is not of great concern, two to three parameters should be sufficient. We did not encounter any computational difficulties when fitting the model. The fit is generally very good, with variations by cause.

A subset of the regression output, namely the parameter estimates and accompanying standard errors for causes 1–3, are presented in Table 3. Significance levels are provided, where '***' indicates a p-value less than 0.001, '**', less than 0.01, '*', less than 0.05, and '.', less than 0.1.

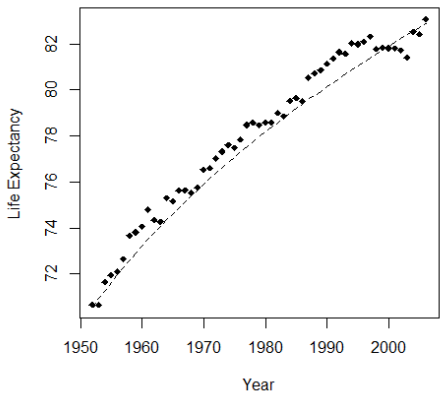
Table 3: Regression Parameter Estimates and Standard Errors for Infectious and Parasitic Diseases (1), Cancer (2), Diseases of the Circulatory System (3)

		Cause 1			Cause 2			Cause 3		
Parameter		Estimate	Standard Error		Estimate	Standard Error		Estimate	Standard Error	
Intercept		-6.4984	0.0177	***	-9.1892	0.0506	***	-8.5390	0.0470	***
age	1–4	-1.5889	0.0278	***	0.3548	0.0554	***	-1.6305	0.0644	***
	5–9	-3.4043	0.0450	***	-0.2892	0.0570	***	-2.4341	0.0788	***
	10–14	-3.9215	0.0564	***	-0.3441	0.0580	***	-1.7452	0.0718	***
	15–19	-3.1949	0.0460	***	-0.1535	0.0571	**	-1.5202	0.0624	***
	20–14	-2.5116	0.0354	***	-0.0614	0.0560		-1.0342	0.0565	***
	25–29	-1.7558	0.0303	***	0.3765	0.0542	***	-0.6157	0.0543	***
	30–34	-1.5628	0.0279	***	0.9182	0.0527	***	-0.2468	0.0516	***
	35–39	-1.6659	0.0279	***	1.5095	0.0518	***	0.1214	0.0506	*
	40–44	-1.5964	0.0272	***	2.0693	0.0513	***	0.5973	0.0493	***
	45–49	-1.5356	0.0261	***	2.5737	0.0510	***	1.1151	0.0484	***
	50–54	-1.6035	0.0250	***	2.9887	0.0509	***	1.6657	0.0479	***
	55–59	-1.5107	0.0243	***	3.3335	0.0508	***	2.2367	0.0475	***
	60–64	-1.3470	0.0237	***	3.6493	0.0508	***	2.8685	0.0473	***
	65–69	-1.0950	0.0224	***	3.9960	0.0508	***	3.5441	0.0472	***
	70–74	-0.7877	0.0216	***	4.3386	0.0507	***	4.2140	0.0471	***
	75–79	-0.5407	0.0213	***	4.6693	0.0507	***	4.8593	0.0471	***
	80–84	-0.3877	0.0220	***	4.9616	0.0508	***	5.4090	0.0471	***
	85+	-0.1546	0.0222	***	5.2540	0.0508	***	5.9694	0.0471	***
t		-0.0619	0.0009	***	-0.0188	0.0017	***	-0.0549	0.0022	***
t*age	1–4	-0.0152	0.0015	***	-0.0088	0.0019	***	0.0270	0.0028	***
	5–9	0.0105	0.0021	***	0.0022	0.0019		0.0195	0.0033	***
	10–14	0.0197	0.0023	***	0.0014	0.0020		0.0017	0.0033	
	15–19	0.0043	0.0021	*	0.0018	0.0019		0.0274	0.0026	***
	20–14	0.0003	0.0017		0.0084	0.0019	***	0.0311	0.0025	***
	25–29	-0.0292	0.0017	***	0.0079	0.0018	***	0.0208	0.0024	***
	30–34	-0.0217	0.0015	***	0.0081	0.0018	***	0.0295	0.0023	***
	35–39	0.0001	0.0013		0.0117	0.0018	***	0.0296	0.0023	***
	40–44	-0.0038	0.0013	**	0.0119	0.0018	***	0.0265	0.0023	***
	45–49	0.0006	0.0013		0.0132	0.0017	***	0.0260	0.0022	***
	50–54	0.0188	0.0011	***	0.0124	0.0017	***	0.0230	0.0022	***
	55–59	0.0257	0.0011	***	0.0132	0.0017	***	0.0217	0.0022	***
	60–64	0.0292	0.0011	***	0.0143	0.0017	***	0.0233	0.0022	***
	65–69	0.0389	0.0010	***	0.0132	0.0017	***	0.0215	0.0022	***
	70–74	0.0426	0.0010	***	0.0132	0.0017	***	0.0239	0.0022	***
	75–79	0.0515	0.0009	***	0.0136	0.0017	***	0.0260	0.0022	***
	80–84	0.0611	0.0009	***	0.0150	0.0017	***	0.0323	0.0022	***
	85+	0.0736	0.0009	***	0.0166	0.0017	***	0.0415	0.0022	***

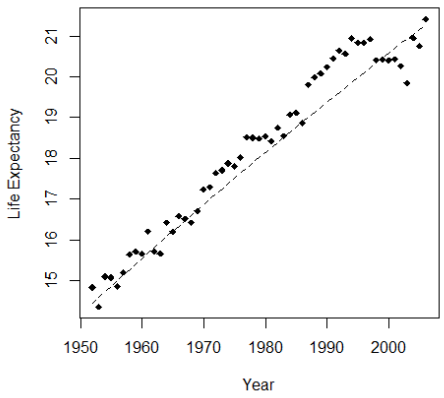
Figures 2a and 2b present the (residual) life expectancy at birth and at retirement age, respectively, for women in France. The observed life expectancy is plotted with points, the fitted life expectancy with dashed lines. As a result of our

Figure 2: Life Expectancy for Women in France

(a) At birth



(b) At retirement age (65)



model selection criteria, the life expectancy fit is good. The observed life expectancy appears to be decaying, which the fit is able to capture.

4.3 Causal Mortality Shocks

Figures 3a and 3b present the impact of eliminating cancer (cause 2) on life expectancy at birth and retirement age, respectively. Note that cause-elimination has been assumed for all age-groups. The cure for cancer results in a life expectancy gain of 3.38 years at birth and 2.15 years at age 65 in 2008. These figures are comparable to results found for women in the US; see Kaishev *et al.* (2007), Tables 2 and 3. In the case of independence, they obtain a life expectancy gain of 3.34 years at birth and 1.97 years at age 65 with a Gaussian copula and 3.46 and 2.16 years, respectively, with a Student's t-copula.

It is evident that the *hypothetical* gain in life expectancy from eliminating cancer is larger in more recent calendar years; most especially for older ages as demonstrated in Figure 3b. The importance of cancer as a cause of death has been increasing with time and is most relevant for older adults. This is intuitive, but difficult to discern from observed data only, such as plots provided in Figure 1.

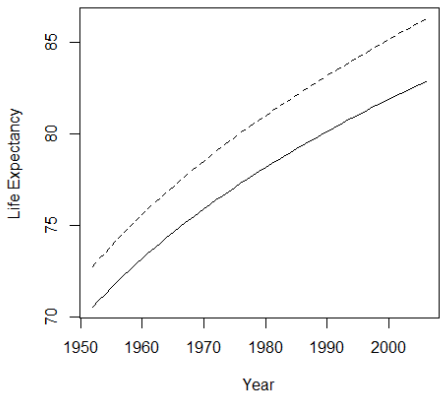
For example, one might perceive a decrease in cancer deaths for a specific age-group. To gain insight into the behaviour of cancer mortality, it should be considered in relation to total mortality. It is plausible that an age-group is transitioning to better overall mortality, but that cancer prevalence is increasing as a cause of death, rather than decreasing. Finally, the age-groups must be aggregated to obtain the impact of cancer on life expectancy over time. This is a difficult obstacle that the multinomial model overcomes.

4.4 Forecasting Residual Life Expectancy

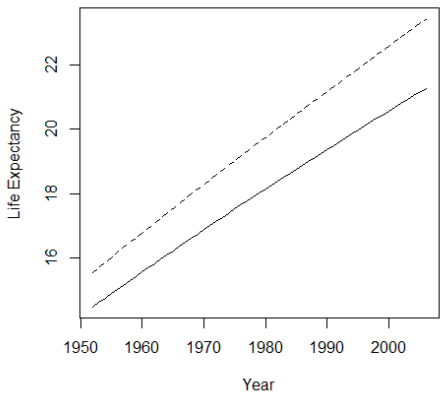
Time is treated as a continuous covariate in the model. This has the benefit of avoiding time-series analysis for forecasting purposes. However, as with any form of forecasting, the implications of projections must be carefully considered. Consequently, we limit the forecasting period to a ten year horizon.

Figure 3: The Impact of Eliminating Cancer on Fitted Life Expectancy

(a) At birth



(b) At retirement age (65)



For the forecasting period, we emphasize the uncertainty driven by potential causal shocks rather than those originating from the process and estimated parameters. A *crude* idea of uncertainty is provided by comparing the forecasted life expectancy under the scenario that cause i is eliminated for each i over the entire age-range. Figure 4 presents the fitted and forecasted life expectancy, where the forecast labelled i represents the scenario that cause i is eliminated. For example, the scenario of a cure for cancer is represented by forecast 2, which has a very large impact on life expectancy at birth as well as life expectancy at retirement age. The projection labelled 0 represents the scenario of no causal shocks. Deaths from cancer (cause 2) and the circulatory system (cause 3) are especially relevant, which is evident in Figure 4 by the magnitude and sustainability of the increase in life expectancy. Diseases of the circulatory system were the most important causes of death about 50 years ago. While cancer is already the most important cause over the entire age-range (see Figure 4a), it is expected to become the most prevalent one at older ages within the next five years (Figure 4b). Deaths from the remaining causes (causes 1, 4, 5) are less than those from cancer and the circulatory system and therefore less relevant to life expectancy.

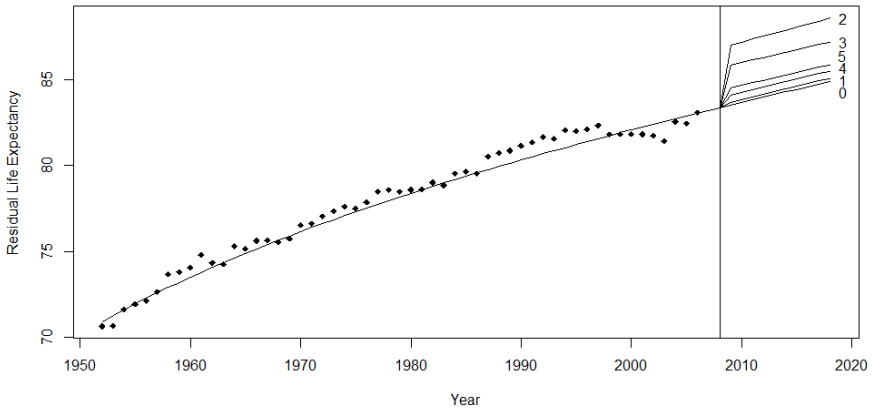
5 Conclusions

The aim of this paper is to provide an alternative approach to cause-of-death mortality modelling. This is especially relevant under current European pension reforms and developments in solvency 2 regulations. Previous work applied in practice has mainly considered modelling causal forces of mortality. A consequence of the instantaneous perspective is that survival is treated differently from death. In the multinomial logistic framework that utilizes annual probabilities, survival is a competing outcome and is treated the same way as the other outcomes. The annual approach assigns less probability to survival as a result of cause-elimination than does the instantaneous approach. Without prior knowledge of the governing behaviour between the various outcomes, we adjust all remaining outcomes similarly; that is, proportional to their probability.

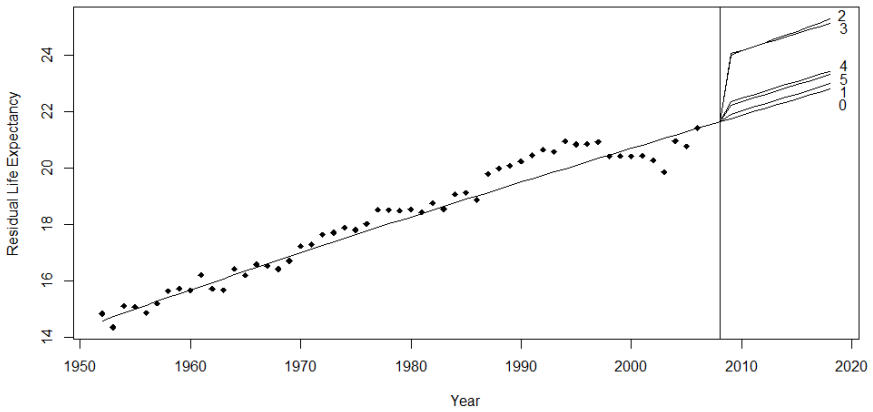
The multinomial logistic framework is easy to implement. It is also easy to

Figure 4: Forecasted Life Expectancy conjoined with Cause-Elimination

(a) At birth



(b) At retirement age (65)



quantify the impact of cause-elimination or shocks on mortality metrics such as life expectancy, since the model provides an intuitive framework for any combination of shocks on the various considered causes. Given the accessibility of this modelling framework, it can readily be used in practice and broaden the perspective offered by currently used methods. Finally, the framework allows for a straightforward implementation of information with respect to known links between the various causes; although such links are not investigated in this paper.

Treating time as a continuous covariate is appealing since it avoids time-series analysis for forecasting purposes, making projections a trivial exercise. However, as with any form of forecasting, the implications of projections must be carefully evaluated. Thus, a shift from continuous to categorical time is worthy of exploration, and must be carefully considered to avoid violating the law of parsimony.

Acknowledgements

The authors would like to acknowledge the financial support of ARC Linkage Grant Project LP0883398 Managing Risk with Insurance and Superannuation as Individuals Age with industry partners PwC, APRA and the World Bank as well as the support of the Australian Research Council Centre of Excellence in Population Ageing Research (project number CE110001029).

References

Anderson, R. N. (1999). US decennial life tables: 1989-91. United States life tables eliminating certain causes of death. *DHHS Publication No (PHS) 99-1150-4*, **1**(4).

Bayo, F. (1968). Life tables: 1959-61. United States life tables by causes of death: 1959-61. *Public Health Service Publication No 1252*, **1**(6).

Booth, H. and Tickle, L. (2008). Mortality modelling and forecasting: A review of methods. *Annals of Actuarial Science*, **3**, 3–43.

Booth, H., Maindonald, J., and Smith, L. (2001). Age-time interactions in mortality projection: Applying Lee-Carter to Australia. *Working Papers in Demography*, **85**, 2–28.

Borooah, V. K. (2002). *Logit and Probit: Ordered and Multinominal Models.* Thousand Oaks, CA: Sage Publications.

Bowers, N., Gerber, H., Hickman, J., Jones, D., and Nesbitt, C. (1986). *Actuarial Mathematics*. Society of Actuaries.

Bradshaw, D., Groenewald, P., Laubscher, R., Nannan, N., Nojilana, B., Norman, R., Pieterse, D., and Schneider, M. (2003). Initial burden of disease estimates for South Africa, 2000. Technical report, Burden of Disease Research Unit.

Cairns, A. J. G., Blake, D., Dowd, K., Coughlan, G. D., Epstein, D., and Khalaf-Allah, M. (2011). Mortality density forecasts: an analysis of six stochastic mortality models. *Insurance: Mathematics and Economics*, **48**(3), 335–367.

Carriere, J. F. (1994). Dependent decrement theory. *Transactions, Society of Actuaries*, **46**, 45–65.

Caselli, G. (1996). Future longevity among the elderly. In G. Caselli and A. D. Lopez, editors, *Health and Mortality among Elderly Populations*, pages 235–265. Clarendon Press Oxford.

Caselli, G., Vallin, J., and Marsili, M. (2006). How useful are the causes of death when extrapolating mortality trends. An update. *Social Insurance Studies from the Swedish Social Insurance*, **4**.

Chiang, C. L. (1968). *Introduction to Stochastic Process in Biostatistics*. John Wiley and Sons, New York.

Chiang, C. L. (1984). *The Life Table and its Applications*. Malabar: Robert E Krieger Publishing Company.

Curtin, L. R. and Armstrong, R. J. (1988). US decennial life tables: 1979-81. United States life tables eliminating certain causes of death. *DHHS Publication No (PHS) 88-1150-2*, **1**(2).

Dimitrova, D. S., Haberman, S., and Kaishev, V. K. (2013). Dependent competing risks: Cause elimination and its impact on survival. *Insurance: Mathematics and Economics*, **53**(2), 464–477.

Eberstein, I. W., Nam, C. B., and Hummer, R. A. (1990). Infant mortality by cause of death: Main and interaction effects. *Demography*, **27**(3), 413–430.

Elandt-Johnson, R. C. (1976). Conditional failure time distributions under competing risk theory with dependent failure times and proportional hazard rates. *Scandinavian Actuarial Journal*, **1976**(1), 37–51.

Foreman, K. J., Lozano, R., Lopez, A. D., and Murray, C. J. (2012). Modeling causes of death: an integrated approach using CODEm. *Population Health Metrics*, **10**(1).

Gaille, S. and Sherris, M. (2011). Modeling Mortality with Common Stochastic Long-Run Trends. *The Geneva Papers on Risk and Insurance - Issues and Practice*, **36**(4), 595–621.

Greville, T. N. E., Bayo, F., and Foster, R. S. (1975). Life tables: 1969-71. United States life tables by causes of death: 1969-71. *DHEW Publication No (HRA) 75-1150*, **1**(5).

Gutterman, S. and Vanderhoof, I. T. (1998). Forecasting changes in mortality: A search for a law of causes and effects. *North American Actuarial Journal*, **2**(4), 135–138.

Haberman, S. and Renshaw, A. E. (2011). A comparative study of parametric mortality projection models. *Insurance: Mathematics and Economics*, **48**(1), 35–55.

Heligman, L. and Pollard, J. H. (1980). The age pattern of mortality. *J. Institute of Actuaries*, **107**, 49–80.

Hougaard, P. (1984). Life table methods for heterogeneous populations: Distributions describing the heterogeneity. *Biometrika*, **71**(1), 75–83.

Human Mortality Database (2012). University of California, Berkeley (USA), and Max Planck Institute for Demographic Research (Germany). Available at `www.mortality.org` or `www.humanmortality.de`.

Johnson, H. L., Liu, L., Fischer-Walker, C., and Black, R. E. (2010). Estimating the distribution of causes of death among children age 1-59 months in high-mortality countries with incomplete death certification. *International Journal of Epidemiology*, **39**, 1103–1114.

Kaishev, V. K., Dimitrova, D. S., and Haberman, S. (2007). Modelling the joint distribution of competing risks survival times using copula functions. *Insurance: Mathematics and Economics*, **41**(3), 339–361.

Keyfitz, N. (1977). What difference would it make if cancer were eradicated? An examination of the Taeuber Paradox. *Demography*, **14**(4), 411–418.

Lawn, J. E., Wilczynska-Ketende, K., and Cousens, S. N. (2006). Estimating the causes of 4 million neonatal deaths in the year 2000. *International Journal of Epidemiology*, **35**, 706–718.

Lee, R. D. and Carter, L. R. (1992). Modeling and forecasting U.S. mortality. *J. American Statistical Association*, **87**, 659–671.

Liu, L., Johnson, H. L., Cousens, S., Perin, J., Scott, S., Lawn, J. E., Rudan, I., Campbell, H., Cibulskis, R., Li, M., Mathers, C., and Black, R. E. (2012). Global, regional, and national causes of child mortality: an updated systematic analysis for 2010 with time trends since 2000. *The Lancet*, **379**, 2151–2161.

LIWMPC Longevity Research Group (2010). LIWMPC Longevity Research Group Update 2010. *The Institute of Actuaries Australia*.

Lo, S. and Wilke, R. A. (2010). A copula model for dependent competing risks. *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, **59**(2), 359–376.

Manton, K. G. (1986). Past and future life expectancy increases at later ages: Their implications for the linkage of morbidity, disability, and mortality. *Journal of Gerontology*, **41**(5), 672–681.

Manton, K. G. (1991). The dynamics of population aging: Demography and policy analysis. *The Milbank Quaterly*, **69**(2), 309–338.

Manton, K. G. and Myers, G. C. (1987). Recent trends in multiple-caused mortality 1968 to 1982: Age and cohort components. *Population Research and Policy Review*, **6**, 161–176.

Manton, K. G. and Poss, S. S. (1979). Effects of dependency among causes of death for cause elimination life table strategies. *Demography*, **16**(2), 313–327.

Manton, K. G., Tolley, H. D., and Poss, S. S. (1976). Life table techniques for multiple-cause mortality. *Demography*, **13**(4), 541–564.

Manton, K. G., Stallard, E., and Poss, S. S. (1980a). Estimates of U.S. multiple cause life tables. *Demography*, **17**(1), 85–102.

Manton, K. G., Patrick, C. H., and Stallard, E. (1980b). Mortality model based on delays in progression of chronic diseases: Alternative to cause elimination model. *Public Health Report*, **95**(6), 580–588.

Manton, K. G., Stallard, E., and Vaupel, J. W. (1986). Alternative models for the heterogeneity of mortality risks among the aged. *Journal of the American Statistical Association*, **81**(395), 635–644.

McNown, R. and Rogers, A. (1992). Forecasting cause-specific mortality using time series methods. *International Journal of Forecasting*, **8**, 413–432.

Menard, S. (2002). *Applied Logistic Regression Analysis*. Thousand Oaks, CA: Sage Publications.

Murray, C. J., Kulkarni, S. C., and Ezzati, M. (2006). Understanding the coronary heart disease versus total cardiovascular mortality paradox: A method to enhance the comparability of cardiovascular death statistics in the United States. *Circulation*.

Olshansky, S. J. (1987). Simultaneous/muliple cause-delay (SIMCAD): An epidemiological approach to projecting mortality. *Journal of Gerontology*, **42**(4), 358–365.

Olshansky, S. J. (1988). On forecasting mortality. *The Milbank Quaterly*, **66**(3), 482–530.

Park, Y., Choi, J. W., and Lee, D.-H. (2006). A parametric approach for measuring the effect of the 10th revision of the international classification of diseases. *Journal of the Royal Statistical Society. Series C (Applied Statistics)*, **55**(5), 677–697.

Pitacco, E., Denuit, M., Haberman, S., and Olivieri, A. (2009). *Modelling Longevity Dynamics for Pensions and Annuity Business*. Oxford University Press.

Prentice, R. L., Kalbfleisch, J. D., Peterson, A. V., Flournoy, N., Farewell, V. T., and Breslow, N. E. (1978). The analysis of failure times in the presence of competing risks. *Biometrics*, **34**, 541–554.

Richards, S. J. (2009). Selected issues in modelling mortality by cause and in small populations. *British Actuarial Journal*, **15**, 267–283.

Rogers, A. and Gard, K. (1991). Applications of the Heligman/Pollard model mortality schedule. *Population Bulletin of the United Nations*, **30**, 79–105.

Rosén, M. (2006). Forecasting life expectancy and mortality in Sweden–some comments on methodological problems and potential approaches. Technical Report 4, Social Insurance Studies from the Swedish Social Insurance.

Shahraz, S., Bhalla, K., Lozano, R., Bartels, D., and Murray, C. J. L. (2012). Improving the quality of road injury statistics by using regression models to redistribute ill-defined events. *Injury Prevention*.

Tabeau, E., Ekamper, P., Huisman, C., and Bosch, A. (1999). Improving overall mortality forecasts by analysing cause-of-death, period and cohort effects in trends. *European Journal of Population*, **15**, 153–183.

Tabeau, E., Van Den Bergh Jeths, A., and Heathcote, C. (2001). *Forecasting Mortality in Developed Countries. Insights from a Statistical, Demographic and Epidemiological Perspective*. Kluwer Academic Publishers, Dordrecht.

Tsai, S. P., Lee, E. S., and Hardy, R. J. (1978). The effect of a reduction in leading causes of death: Potential gains in life expectancy. *American Journal of Public Health*, **68**(10), 966–971.

Tuljapurkar, S. (1998). Forecasting mortality change: Questions and assumptions. *North American Actuarial Journal*, **2**(4), 127–134.

Vaupel, J. W. and Yashin, A. I. (1983). The deviant dynamics of death in heterogeneous populations. Technical Report RR-83-001, International Institute for Applied Systems Analysis (IIASA).

Wilmoth, J. R. (1995). Are mortality projections always more pessimistic when disaggregated by cause of death? *Mathematical Population Studies*, **5**(4), 293–319.

Wilmoth, J. R. (1996). Mortality projections for Japan: A comparison of four methods. In G. Caselli and A. D. Lopez, editors, *Health and Mortality among Elderly Populations*, pages 266–287. Clarendon Press Oxford.

Wong-Fupuy, C. and Haberman, S. (2004). Projecting mortality trends: Recent developments in the United Kingdom and the United States. *North American Actuarial Journal*, **8**(2), 56–83.

World Health Organization (2012). WHO Mortality Database. `http://www.who.int/whosis/mort/download/en/index.html`.

A A Comparison of Annual and Instantaneous Mortality

We prove Inequality (5) from Section 3.2 by using Newton's generalized binomial theorem. For $0 < a, b < 1$, we have

$$\begin{aligned}(1-b)^a &= 1 - ab + \frac{a(a-1)}{2}b^2 - \frac{a(a-1)(a-2)}{3\cdot 2}b^3 + \ldots,\\ (1-a)^b &= 1 - ab + \frac{b(b-1)}{2}a^2 - \frac{b(b-1)(b-2)}{3\cdot 2}a^3 + \ldots,\end{aligned}$$

such that

$$(1-b)^a-(1-a)^b = \frac{a(a-1)}{2}b^2-\frac{b(b-1)}{2}a^2-\frac{a(a-1)(a-2)}{3\cdot 2}b^3+\frac{b(b-1)(b-2)}{3\cdot 2}a^3+\dots$$

Each pair on the right hand side is positive for $0 < b < a < 1$. That is,

$$a(a-1)\cdots(a-k)b^{k+1}(-1)^{k+1} > b(b-1)\cdots(b-k)a^{k+1}(-1)^{k+1}, \quad k \in \mathbb{Z}_+.$$

To show this we note that $0 < b < a < 1$ and $0 < (k-a) < (k-b)$ for $k \in \mathbb{Z}_+$.

$$\begin{aligned} b < a \quad & \Rightarrow b^k < a^k \\ & \Rightarrow (1-a)b^k < (1-b)a^k \\ & \Rightarrow (1-a)\cdots(k-a)b^k < (1-b)\cdots(k-b)a^k \\ & \Rightarrow (a-1)\cdots(a-k)b^k(-1)^k < (b-1)\cdots(b-k)a^k(-1)^k \\ & \Rightarrow (a-1)\cdots(a-k)b^k(-1)^{k+1} > (b-1)\cdots(b-k)a^k(-1)^{k+1} \\ & \Rightarrow a(a-1)\cdots(a-k)b^{k+1}(-1)^{k+1} > b(b-1)\cdots(b-k)a^{k+1}(-1)^{k+1}. \end{aligned}$$

We obtain the following inequality:

$$(1-b)^a > (1-a)^b, \qquad 0 < b < a < 1.$$

Inequality (5) is proved by taking $a = q(x,t)$, $b = q_j(x,t)$, and noting that $0 < q_j(x,t) < q(x,t) < 1$.